

Εθνικό Μετσόβιο Πολυτεχνείο

Συστήματα και Τεχνολογίες Γνώσης

Μηχανική Μετάφραση Αριθμητικών



Ηλίας Φωτόπουλος
03109106

1 Στόχος

Στόχος της άσκησης αποτελεί η σχεδίαση ενός συστήματος μηχανικής μετάφρασης, το οποίο θα δέχεται ως είσοδο έναν αριθμό από το 0 έως και το 999 και θα παράγει ως έξοδο τον αριθμό σε ολογραφική μορφή. Η μετάφραση γίνεται από τα ελληνικά στα αγγλικά και αντιστρόφως.

2 Υλοποίηση

Ο μηχανικός μεταφραστής χρησιμοποιεί τα μαθηματικά ως ενδιάμεσο επίπεδο μεταξύ των δύο γλωσσών (interlingua). Ουσιαστικά μετατρέπει ένα αλφαριθμητικό στην αριθμητική του απεικόνιση και στην συνέχεια μεταφράζει την αριθμητική απεικόνιση στην επιθυμητή γλώσσα. Αναλυτικότερα η παραπάνω διαδικασία επιτυγχάνεται μέσω των παρακάτω μηχανισμών/συναρτήσεων:

2.1 Συναρτήσεις `elnum/2` και `ennum/2`

Οι συναρτήσεις `elnum/2` και `ennum/2` χρησιμοποιούνται για να χωρίσουν τον ακέραιο σε εκατοντάδες δεκάδες και μονάδες. Αρχικώς κάθε συνάρτηση αντιμετωπίζει τις ειδικές περιπτώσεις αριθμών κάθε γλώσσας (για τα ελληνικά το 11 και 12 όπως φαίνεται παρακάτω):

```
1 elnum(11,[[11],11]):- !.  
2 elnum(12,[[12],12]):- !.
```

Ενώ σε περίπτωση μη ειδικών αριθμών ακολουθείτε η διαδικασία που φαίνεται παρακάτω:

```
1 % Rule for X < 1000  
2 elnum(X,[[N,Y,Z],X]) :-  
3     integer(X), X < 1000,  
4     N is floor(X / 100) * ←  
5     100, % N = hundreds  
6     Y is X mod 100 - X mod 10, ←  
7     % Y = tens  
8     Z is X mod 10. % Z = ←  
9     units
```

Ανάλογη είναι και η διαδικασία που ακολουθεί η συνάρτηση `ennum/2` η οποία έχει όμως διαφορετικούς ειδικούς αριθμούς.

2.2 DCG Rules

Ο επόμενος μηχανισμός στον οποίο στηρίζεται ο μηχανικός μεταφραστής είναι οι DCG Rules. Για τα ελληνικά οι κανόνες αυτοί φαίνονται παρακάτω:

```
1 % 9 Rules (Cases)
2
3 % Special cases 11,12 | Hundreds 100,200 | Tens 10,20 | ↔
  Units 1,2
4 el(N) --> eidikoiarithmoi(N);ekatontades(N);dekades(N);↔
  monades(N).
5
6 % Example: 128
7 el(N) --> ekatontades(N1),dekades(N2),monades(N3),
8 {N is N1 + N2 + N3,N =\= 11, N =\= 12}.
9
10 % Example: 42
11 el(N) --> dekades(N1),monades(N2),
12 {N is N1 + N2,N =\= 11, N =\= 12}.
13
14 % Example: 140
15 el(N)--> ekatontades(N1),dekades(N2),
16 {N is N1 + N2,N =\= 11, N =\= 12}.
17
18 % Example: 302
19 el(N)--> ekatontades(N1),monades(N2),
20 {N is N1 + N2,N =\= 11, N =\= 12}.
21
22 % Example 112
23 el(N)--> ekatontades(N1),eidikoiarithmoi(N2),
24 {N is N1 + N2,N2 =\= 100, N2 =\= 0},!.
```

Ανάλογοι είναι και οι κανόνες για την Αγγλική γλώσσα.

2.3 Λεξικό (Βιβλιοθήκη DCG)

Για να επιτευχθεί η μηχανική μετάφραση χρειαζόμαστε και τα αντίστοιχα λεξικά για την Αγγλική και Ελληνική γλώσσα. Ενδεικτικά ένα μέρος του ελληνικού λεξικού φαίνεται παρακάτω:

```
1 mhden(0) --> μηδέν [].
2 ena(1) --> ένα [].
```

```

3 | duo(2) --> δύο [].
4 | tria(3) --> τρία [].
5 | tessera(4) --> τέσσερα [].

```

2.4 Τελικές Συναρτήσεις

Οι τελικές συναρτήσεις που συνδυάζουν όλα τα παραπάνω και πετυχαίνουν την τελική μετάφραση φαίνονται παρακάτω:

```

1 | % Main Functions
2 |
3 | % Number to Greek
4 | greeknum(X,Y) :-
5 |     elnum(X,[_,Z]),
6 |     el(Z,Y,[ ]).
7 |
8 | % Number to English
9 | englishnum(X,Y) :-
10 |     ennum(X,[_,Z]),
11 |     en(Z,Y,[ ]).
12 |
13 | % Greek to Number
14 | greekalpha(Y,X) :-
15 |     el(Z,Y,[ ]),
16 |     elnum(X,[_,Z]).
17 |
18 |
19 | % English to Number
20 | englishalpha(Y,X) :-
21 |     en(Z,Y,[ ]),
22 |     ennum(X,[_,Z]).
23 |
24 | % Translate from Greek to English
25 | translategreek(Y,Z) :-
26 |     greekalpha(Y,X),
27 |     englishnum(X,Z),
28 |     write('Το νούμερο'), write(Y), write(' στα ←
29 |         αγγλικά είναι : ' ).
30 |
31 | % Translate from English to Greek
32 | translateenglish(Y,Z) :-
33 |     englishalpha(Y,X),

```

```

33 |         greeknum(X,Z) ,
34 |         write('The number' ) , write(Y) , write('in ←
      |         Greek is:') .

```

Το επιθυμητό αποτέλεσμα επιτυγχάνεται μέσω των συναρτήσεων `translategreek/2` και `translateenglish/2`. Όπως περιγράψαμε αρχικώς η συνάρτηση `translategreek/2` σε πρώτο στάδιο μεταφράζει τα ελληνικά σε αριθμητική μορφή (*interlingua* στάδιο) και στην συνέχεια μεταφράζει τον αριθμό στα Αγγλικά στην ολογραφική του μορφή. Ενώ η συνάρτηση `translateenglish/2` δουλεύει ακριβώς με τον ίδιο τρόπο απλά επιτελεί την αντίστροφη διαδικασία.

3 Αμφισημία

Κατά την ανάπτυξη του μηχανικού μεταφραστή υπήρξαν και προβλήματα αμφισημίας όπως τα παρακάτω.

- Στα αγγλικά χρειάζονται δύο λέξεις για την περιγραφή των εκατοντάδων (*one hundred* - *εκάτο*)
- Στα ελληνικά είχαμε τις ειδικές περιπτώσεις του 11 και 12, όπου δεν μπορούσε να ισχύσει ο γενικός κανόνας 11 --> *Δέκα ένα*
- Όπως αναφέρεται και στην εκφώνηση έχουμε πρόβλημα για το 100 (*εκατό* ή *εκατόν*), το οποίο λύθηκε με ειδικό χειρισμό και στο λεξικό και στους κανόνες DCG.